



ASSIMILATION OF DISCHARGE DATA INTO SEMI-DISTRIBUTED CATCHMENT MODELS FOR SHORT TERM FLOW FORECASTING: CASE STUDY OF THE SEINE RIVER BASIN

Simon Munier, Xavier Litrico, Gilles Belaud, Charles Perrin

► To cite this version:

Simon Munier, Xavier Litrico, Gilles Belaud, Charles Perrin. ASSIMILATION OF DISCHARGE DATA INTO SEMI-DISTRIBUTED CATCHMENT MODELS FOR SHORT TERM FLOW FORECASTING: CASE STUDY OF THE SEINE RIVER BASIN. *Journal of Hydrologic Engineering*, 2014, 20, pp.05014021. 10.1061/(ASCE)HE.1943-5584.0001054 . hal-01162521

HAL Id: hal-01162521

<https://hal.science/hal-01162521>

Submitted on 11 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ASSIMILATION OF DISCHARGE DATA INTO SEMI-DISTRIBUTED CATCHMENT MODELS FOR SHORT TERM FLOW FORECASTING: CASE STUDY OF THE SEINE RIVER BASIN

Simon Munier¹, Not a member, ASCE

Xavier Litrico², Member, ASCE

Gilles Belaud³, Not a member, ASCE

and Charles Perrin⁴, Not a member, ASCE

ABSTRACT

This study addresses the sensitivity of short-term flow forecasting in the Seine River basin (43,800 km², France) to the spatial distribution using a semi-distributed model (*Transfer with GR*, TGR). The basin was decomposed into intermediate basins depending on the gauging stations selected for this study. A lumped hydrological model was applied on each intermediate basin and a routing model was used to propagate the discharge through the river network. Discharge data at the gauging stations were assimilated using a Kalman filter and tests for flow forecasting were performed with a lead time up to 72 h. Several spatial configurations, defined by a selection of one or several gauging stations, were tested and the performances were compared to a reference lumped model currently used operationally by the regional flood forecasting centre. Results showed that the forecasting performance improves with an increase in the degree of spatialization. Nevertheless this improvement was not systematic and the integration of some particular gauging stations degraded the model

¹Estellus, 93 boulevard de Sébastopol, 75002 Paris, France. simon.munier@gmail.com.

²LyRE, R&D center of Lyonnaise des Eaux, 91 rue Paulin, 33029 Bordeaux, France. xavier.litrico@lyonnaise-des-eaux.fr.

³Montpellier SupAgro, UMR G-EAU, 2 place Pierre Viala, 34060 Montpellier Cedex 1, France. belaud@supagro.inra.fr.

⁴Irstea, UR HBAN, 1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony Cedex, France. charles.perrin@irstea.fr.

performance. In addition, it was shown that integrating some other stations (generally the most upstream) led to a negligible improvement. This suggests that in an operational context, where the model has to be robust and computationally efficient, some efforts should focus on finding the optimal spatial distribution, which is not necessarily the one using all the available stations.

Keywords: flow forecasting, spatial distribution, semi-distributed model, hydrology, routing model, lumped model, Seine River

INTRODUCTION

Flood forecasting remains a difficult issue for hydrologists (see e.g., Young 2002; Todini 2007; Liu et al. 2012). In spite of a variety of available models and tools (see Cloke and Pappenberger 2009, and references therein), the improvements in flood forecasting tools are slow and there is large margin of progress (Kealey 2007; Welles et al. 2007). Namely, the role of spatial distribution on model efficiency remains a matter of debate in the hydrological community. In a flood forecasting context, the sensitivity of model forecasts quality seems dependent on various physical factors and sources of uncertainty. The lumped approach, though simple, find limitations for events showing a large spatial variability (see e.g., Cole and Moore 2009). Similarly, fully distributed models do not appear to be the panacea given their complexity and their lack of overall superiority (Smith et al. 2012). Hence semi-distributed approaches are often considered as a good trade-off between complexity and efficiency (see e.g., Amengual et al. 2008). Indeed, they couple rainfall-runoff models on sub-catchments (here defined by gauging stations) and simple propagation tools like unit hydrographs or lag-and-route methods (see e.g., Lerat et al. 2012). In a forecasting context, one key issue is to simultaneously assimilate the information available in real-time. This may be spatially distributed data like snow cover in mountainous regions (see e.g., Nester et al. 2012) or more classically the flow observations at the gauging stations within the catchment (Mendoza et al. 2012).

This study focuses on a large catchment in France, the Seine River upper basin at Paris

(43,800 km²), where there are major socio-economic potential impacts in the capital city of Paris. The regional flood forecasting centre (FFC) based in Paris is in charge of routinely issuing forecasts on the Seine and its tributaries to feed the national flood warning map (www.vigicrues.gouv.fr). In flood conditions, the exceedance of warning thresholds is forecasted to provide information for civilian security services, which ask for three-day anticipation at Paris to have sufficient time for evacuating people and installing protections against flooding (Lacaze et al. 2011). The FFC has implemented a forecasting system which uses several hydrological and hydraulic models that are fed by hydro-meteorological observations and forecasts received in real-time. Among these models is a conceptual lumped hydrological model which only simulates the discharge at the basin outlet, at which it was calibrated. Due to its lumped structure (i.e. no explicit channel routing is made), it is not able to simulate the discharge at upstream gauging stations and to account for information available at these sites. However, it is likely that the information from these upstream stations spread over the basin can be useful to improve forecasts at the basin outlet.

The objective of the study was to investigate the sensitivity of flood forecasts to the spatial distribution of inputs and outputs using a semi-distributed hydrological model and a Kalman-type real-time assimilation scheme. The Seine River basin was first decomposed into several intermediate basins based on the location of the gauging stations. Then several levels of spatial discretization were considered and the sensitivity of model performance at the basin outlet with increasing spatial resolution was analysed. The semi-distributed model has two main components: a hydrological module based on the lumped model and applied on each intermediate basin, and a flow routing module representing the flow propagation between gauging stations. In order to check whether the semi-distributed model could provide improvement over the lumped model, which is currently used in an operational context by the French FFC, the performance of the semi-distributed model was benchmarked against the performance of the lumped model used as a reference.

The next sections successively present the Seine River basin and the data used, the

forecasting model and the assimilation framework, and the testing methodology. Then results are presented and discussed in section 5.

THE SEINE RIVER UPPER BASIN AND AVAILABLE DATA

The Seine River basin drains an area of 43,800 km² upstream of Paris, which is the outlet considered in this study. From upstream to downstream, three major tributaries contribute to the flow of the Seine River: the Yonne, Loing and Marne Rivers. The rivers mainly flow westwards. The relief is quite low over the basin with altitudes mainly below 500 m except in the upper Yonne basin with altitudes up to 900 m, which makes this tributary the most reactive within the basin (see e.g., Billen et al. 2009; Viennot et al. 2009 for a detailed description of the basin).

Four large dams, with a total storage capacity of about 800 hm³ were built in the upper part of the basin, on the Marne, Seine, Aube and Yonne Rivers, respectively. They are managed to regulate downstream flows, especially for the Paris region, with the two objectives of low-flow augmentation (for domestic, industrial and agricultural water supply) and flood alleviation (Villion 1997). Flooding is a major natural risk in the Paris region, as shown by the consequences of the major 1910 flood. This flood reached 2,400 m³/s, which corresponds to about a 100-year return period. It is estimated that such a flood would today directly impact 850,000 people who live in zones liable to flooding within the Paris region and cause direct damages estimated to more than 10 billion Euros (<http://vertigo.revues.org/14339#ftn1>). Hence the early anticipation of such events may help mitigating their catastrophic consequences.

The basin is under an oceanic climate, with mean precipitations of 790 mm/yr, mean temperature of 11.1°C, and mean potential evapotranspiration (PE) of 690 mm using the Penman formula (average over the 1958-2011 period using the SAFRAN reanalysis, see Vidal et al. 2010). Precipitation is almost evenly distributed within the year. A network of 68 rain gauges were used in this study (see Fig. 1), as well as a network of 10 flow gauging stations distributed over the basin (see details in Table 1). These ten stations

are part of the stations considered by the FFC in their operational forecasting system. Four gauging stations (La Ferté-sous-Jouarre, Bazoches-lès-Bray, Courlon-sur-Yonne and Episy) control a large part of the basin upstream Paris, with a cumulated area of 33,520 km² (i.e. 77% of the basin). They are all situated downstream of the dams and therefore integrate the impact of dam management. Hence they are interesting to consider in a flood forecasting context. Within the Yonne basin, four additional stations (Dornecy, Arcy-sur-Cure, Dissangis and Aisy-sur-Armanon) were considered to better account for the upstream part, which receives the largest cumulated rainfall amounts on the basin. An additional upstream station (Châlon-sur-Marne) was also considered on the Marne River, given the elongated shape of this catchment.

Hourly time series of rainfall, streamflow and PE were available over the 1991-2009 period. Since there are many missing data before 1995, the 1991-1994 period will be used only as a warm-up period. The percentage of missing values for the 1995-2009 is reasonable (see Table 1 for the gauging stations), except for the Bazoches-lès-Bray station for which no data were available before 1999. Large flood events occurred in Paris in the years 1995, 1999, 2000 and 2001. The selected periods are therefore interesting for analysing the behaviour of the model in high flow conditions.

SEMI-DISTRIBUTED HYDROLOGY-ROUTING COUPLED MODEL

This section details the semi-distributed model built in this study to investigate the impact of spatial discretization. We first detail the way the basin was spatially split. Then the structures of the hydrological and flow routing sub-models are described, before presenting the assimilation framework used to run the coupled model in forecasting mode.

The semi-distributed modelling approach

Modelling principle

The coupled model simulates the discharge at each gauging station. Therefore it can assimilate discharge observations at these stations and propagate the improvements due to state corrections downstream. The semi-distributed approach adopted in this study is

intermediate between lumped and fully distributed approaches. It divides the basin into hydraulically connected sub-catchments.

Here, we aim at representing the discharge at the gauging stations on the main streams. However we do not need to know the exact states of the system between these target points. Therefore the intermediate basins were defined between these stations. Here an intermediate basin is defined by an outlet (corresponding to a gauging station) and at least one (or possibly several) upstream catchment (also defined by gauging stations). Therefore the intermediate basin is the area that drains water to the river reach(es) between the downstream and the upstream station(s).

For each intermediate basin, the model represents two types of water transfers (see Figure 2): the hydrological transfer represents the transformation of rainfall over the basin into discharge that is injected into the main stream as lateral inflows, and the hydraulic transfer (flow routing) corresponds to the discharge propagation through the main streams to a downstream station (see also Lerat et al. 2012). On upstream sub-basins, only the hydrological part of the model is applied.

Spatial discretization

Regarding flow propagation in the main streams, the contribution to streamflow due to rainfall on the intermediate basin may be considered as lateral inflows. The distribution of this lateral discharge along the river stretch has an impact on the downstream discharge (Fan and Li 2006; Munier 2009; Lerat et al. 2012). In the following, lateral flows will be decomposed into concentrated or uniformly distributed lateral discharges. As shown in Figure 2, the flow routing module simulates the propagation of upstream and lateral discharges, whereas the hydrological module simulates the total discharge due to rainfall. The output of the hydrological module is injected into the routing module as lateral discharges following an established longitudinal distribution depending on the hydrographic configuration of the intermediate basin. As done by Lerat et al. (2012), we used the drained area curve to determine the spatial distribution of hydrological inputs. This curve is obtained from a Digital

Elevation Model and represents the area drained by the river stretch with respect to the longitudinal abscissa (see Figure 2). It allows to locate the main tributaries which are considered as concentrated lateral inflows. The remaining surface is considered as a uniformly distributed lateral inflow. The output of the hydrological model is then decomposed into one or several lateral inflows (concentrated and distributed) weighted by their respective draining areas. Munier (2009) showed that considering tributaries draining an area less than 20 % of the total intermediate basin area as concentrated lateral inflows does not improve the simulation. The contribution of these tributaries are considered as distributed lateral inflows. This criterion is used to limit the number of concentrated lateral inflows.

Networking

Considering the structure of the model, and especially the routing model, it is very simple to couple multiple models representing multiple intermediate basins. For each intermediate basin, the hydrological module computes the contribution due to rainfall, using rainfall and evapotranspiration data over the intermediate basin. Then the routing module propagates this lateral discharge as well as any observed upstream discharges (discharge at the outlet of upstream basins) to compute the downstream discharge. Figure 3 shows an example configuration: the discharge at station #5 results from the propagation of the discharge at the four upstream stations (#1 to 4) and the flow produced in the intermediate basin (between station #5 and stations #1 to 4); the discharge at the downstream station #6 results from the propagation of the flow at the upstream station #5 and the flow produced on the intermediate basin (between stations #6 and 5).

Model description

Hydrological model

We used a lumped approach representing the rainfall-runoff transformation at the sub-catchment (or intermediate catchment) scale. The hydrological model is derived from the GR (*Génie Rural*) model developed by Berthet et al. (2009). It is defined by a production store, a function representing surface-subsurface exchanges and a routing store. The GRP

model, an extension of the GR model, integrates a data assimilation procedure that allows to correct the model states by assimilating the discharge at the outlet (Berthet et al. 2009). The GRP (*Génie Rural-Prévision*) model is currently used in operational conditions by the FFC for real-time forecasting.

In this study, the original GR model was coupled with a hydraulic linear model. The routing reservoir was linearized. Hence a linear Kalman filter is applied to correct the hydraulic model states as well as the routing reservoir states. The linearised reservoir corresponds to a first order filter like those defined in the flow routing model (see next section). Figure 4 presents the structure of the hydrological model. Three parameters must be calibrated for each intermediate or upstream catchment:

- S (mm): the capacity of the production reservoir,
- IGF (-): the intercatchment groundwater function coefficient,
- K_R (s): the time constant of the routing reservoir.

Inputs of the hydrological model are the spatial average of precipitation P and potential evapotranspiration PE over the intermediate basin. For each intermediate basin, PE is computed from temperature and incoming radiation as done by Oudin et al. (2005) whereas P is obtained from raingauges inside the basin.

Flow routing model

The flow routing model represents the discharge propagation through the river stretch. It accounts for upstream and lateral (concentrated and uniformly distributed) discharges. Many propagation models have been proposed in the literature, such as unit hydrographs or Muskingum type models (e.g., Perumal et al. 2009). Nevertheless, such models are generally not adapted to multiple inflows (unless by adding new parameters) and to data assimilation. The model presented here is the propagation model developed by Munier (2009), hereafter called LLR (standing for Linear Lag-and-Route). It has been chosen for its ability to account for upstream as well as lateral inflows using a small number of parameters. Additionally, it

is particularly well adapted to data assimilation such as the Kalman filter.

The flow routing model is based on the Saint-Venant equations that describe the 1D free surface flow dynamics. These equations, classically used to represent the flow propagation in a river stretch, may be solved analytically after some simplifications. The main interest of using these equations is that they are described by the physical characteristics of the river stretch (geometry, roughness). Here, we used this property to describe the propagation of every input discharges (upstream and lateral) using only two parameters, which limits identifiability problems that may occur with over-parametrised models (see, e.g., Duan et al. 1992).

The Saint-Venant equations are first linearised around a uniform flow and transposed into the frequential domain using the Laplace transform. Under this form, it is possible to solve the equations analytically, leading to a linear transfer function between the upstream and downstream discharges (Munier et al. 2008). Under the assumption of negligible upward waves, transfer functions relating concentrated and uniformly distributed lateral flows to the downstream discharge can be derived.

The river stretch is schematised as in Figure 5. In this section, the following notations are used: t is the time (s), x the abscissa along the river stretch (m), X the length of the river stretch (m), x_{Pi} the abscissa of the i -th concentrated lateral discharge (m), Q_0 the upstream discharge (m^3/s), Q_X the downstream discharge (m^3/s), Q_{Pi} the i -th concentrated lateral discharge (m^3/s), Q_D the uniformly distributed lateral discharge (m^3/s).

The derivation of transfer functions from the linearised Saint-Venant equations transposed into the Laplace domain is given in Appendix S1 and leads to:

$$Q_X(s) = TF_0(s)Q_0(s) + \sum_i TF_P(x_{Pi}, s)Q_{Pi}(s) + TF_D(s)Q_D(s) \quad (1)$$

where s denotes the Laplace variable, TF_0 , TF_P and TF_D the transfer functions related to the upstream discharge and the concentrated and uniformly distributed lateral discharges,

respectively.

To compute the downstream discharge Q_X , Equation 1 has to be transposed back into the time domain. In order to avoid high computation costs due to convolution algorithm, the previous transfer functions are approximated by first order with delay transfer functions using the Moment Matching Method, as done for example by Munier et al. (2008). Such transfer functions have three main advantages: (1) they simulate the delay and attenuation phenomena which characterise flow propagation, (2) they are easily transposed into the time domain since they represent simple Ordinary Differential Equations (ODE) and (3) they are particularly well suited to simple data assimilation techniques due to their linearity. The three approximate transfer functions can then be written as:

$$TF_0(s) \approx \frac{e^{-\tau_0 s}}{1 + K_0 s} \quad (2)$$

$$TF_P(x_{Pi}, s) \approx \frac{e^{-\tau_P(x_{Pi})s}}{1 + K_P(x_{Pi})s} \quad (3)$$

$$TF_D(s) \approx \frac{1}{1 + K_D s} \quad (4)$$

where τ_0 and τ_P are the delay values, K_0 , K_P and K_D the first order time constants. One may note that TF_D has no delay, which can be explained by the fact that the distributed lateral inflow is injected all along the river reach, and then right upstream the downstream end.

Using analytical transformations (Laplace transform and Moment Matching Method), the parameters of the approximate transfer functions can be expressed as functions of the physical characteristics of the river stretch (see Appendix S2). These expressions can be simplified in order to reduce the number of parameters. Indeed, $\tau_P(x_{Pi})$, $K_P(x_{Pi})$ and K_D

can be expressed as functions of x_{Pi}/X , τ_0 and K_0 .

$$\begin{cases} \tau_P(x_{Pi}) &= \left(1 - \frac{x_{Pi}}{X}\right) (\tau_0 + K_0) - \sqrt{1 - \frac{x_{Pi}}{X}} K_0 \\ K_P(x_{Pi}) &= \sqrt{1 - \frac{x_{Pi}}{X}} K_0 \\ K_D &= \frac{\tau_0 + K_0}{2} \end{cases} \quad (5)$$

Note that the delay parameter τ_P may reach negative values. Since such a case is physically unrealistic, the lower bound for τ_P was set to 0 and parameter K_P is computed consequently:

$$\text{If } \tau_P(x_{Pi}) < 0, \text{ then } \tau_P(x_{Pi}) = 0 \text{ and } K_P(x_{Pi}) = \left(1 - \frac{x_{Pi}}{X}\right) K_0 \quad (6)$$

Last, the flow routing model is described by three types of transfer functions related to upstream, concentrated and uniformly distributed lateral discharges (Equations (2-4)). The discretized form of such transfer functions is described in Appendix S3. If the relative positions of the concentrated lateral discharges are known (x_{pi}/X), only two parameter are necessary to define the model: the time delay τ_0 and the first order time constant K_0 related to the transfer of the upstream discharge. As stated earlier, these parameters can be computed analytically from the physical characteristics of the river reach. Nevertheless, simulation results may be improved if they are calibrated, which is done in this study.

The TGR coupled model

For an intermediate basin, the integrated model, named TGR for Transfer with GR (the selected hydrological model), is obtained by coupling the two sub-models (flow routing and hydrology). The flow routing part and the routing reservoir of the hydrology part are linear. This linear part of TGR is named LRK. The remaining part of the hydrological model (production reservoir and exchange function) is named GRK. Figure 6 presents the scheme of the TGR model for an intermediate basin. For each intermediate basin, five parameters must be identified (τ_0 , K_0 , K_R , S and IGF) using P , PE and Q_0 as inputs and Q_X as output. The relative positions x_{Pi}/X are determined from the drained area curve. For

an intermediate basin with several upstream basins, the relative position of each upstream discharge is also known, and the transfer of the upstream discharges is done consequently, as for concentrated lateral inflows. For a sub-catchment basin with no upstream station, only the hydrological model is applied and thus only the parameters related to the transformation of rainfall into downstream discharge are identified. Several networks of possibly multiple intermediate basins are presented in section 5 through the example of the Seine River upper basin.

Data assimilation for flood forecasting

As pointed out by Liu et al. (2012), the manual correction of the model states by human forecasters, based on their expert interpretation of the discrepancies between model simulations and observed discharges, is still widely practiced in operational forecasting. In an operational context, data assimilation techniques allow to use observations received in real-time to automatically correct the model outputs and refine the forecasts at each time-step. Various assimilation techniques proposed in the literature can be applied for hydrological forecasting (see e.g., Refsgaard 1997). Here we chose the widely used Kalman filter (Kalman 1960). Indeed the coupled model presented in the previous section is particularly well adapted to the Kalman filter, because of its linear part and the simplicity of the state equations. The Kalman filter, which is applied on the linear part (LRK) of the TGR model, is briefly described in the following.

Note that another Kalman filter, such as the Ensemble Kalman Filter (EnKF), could have been used in order to correct not only the state variables of LRK, but also the state variables of the non-linear part (GRK). Examples of existing ensemble flood forecasting systems can be found in Cloke and Pappenberger (2009) and in McMillan et al. (2013). However, as pointed out by Cloke and Pappenberger (2009) and Rabuffetti and Barbero (2005), the characterization of meteorological input uncertainties, which have an important impact in ensemble data assimilation procedures, remain a key challenge. In this paper, the effect of including one or several upstream stations is investigated; the main physical

process involved is thus the flow routing between stations. As stated by Young (1974), most of the non linearities come from the hydrological processes, the routing processes being more linear, all the more so as only high flows are considered in the flood forecasting context. Considering the robustness of the linear Kalman filter, it has been preferred here to more sophisticated extensions.

The Kalman filter is applied to a discretized Linear Time-Invariant (LTI) model described by the following equations:

$$x(k) = A x(k-1) + B u(k) + w(k) \quad (7)$$

$$y(k) = C x(k) + D u(k) + v(k) \quad (8)$$

where k is the discretized time, x the state of the system, u the input and y the output. The input vector gathers the outputs of the GRK model applied on every upstream and intermediate basins. The output vector represents the simulated discharges at every gauging stations. In the LRK model, the output only depends on the state, so that the matrix D is null. The random variables w and v represent the process noise and measurement noise, respectively. They are supposed to be independent and described by a normal probability law with covariance matrices Q for w and R for v . Both matrices are then diagonal. In practice, process and measurement noise covariances are hardly quantifiable and are chosen empirically. In this study the measurement noise is supposed to be lower than the process noise, and values of R are consequently lower than those of Q . Besides, different tests have been conducted to estimate the sensitivity of the forecast performances against values of Q (presented in Munier 2009). Results showed that although the hydrological states (those relative to the routing reservoir) are the most uncertain, little impact on the forecast performances has been observed. Here, we chose a value of 1 for the process noise and 0.1 for the measurement noise.

The Kalman filter is described by the following equations:

$$\hat{x}(k|k-1) = A \hat{x}(k-1|k-1) + B u(k) \quad (9)$$

$$\tilde{y}(k) = y(k) - C \hat{x}(k|k-1) \quad (10)$$

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k) \tilde{y}(k) \quad (11)$$

where $\hat{x}(k|k-1)$ and $\hat{x}(k|k)$ are the prediction and the update states at time k , respectively, y the measurements, $\tilde{y}(k)$ the measurement error (or innovation) and $K(k)$ a matrix called the *Kalman Gain*. The corrected output is then given by:

$$\hat{y}(k) = C \hat{x}(k|k) \quad (12)$$

Missing values are replaced by model prediction output given by $C\hat{x}(k|k-1)$. For instance, if the i -th line of vector y is missing, then this line is replaced by the i -th line of the model prediction output vector. This is equivalent to considering that the model is perfect at this station.

The Kalman gain is computed so as to minimise the covariance $P(k|k)$ of the state error $e(k|k) = x(k) - \hat{x}(k|k)$. The optimal Kalman gain is obtained from (e.g., Brown and Hwang 1992):

$$K(k) = P(k|k-1)C^T (CP(k|k-1)C^T + R)^{-1} \quad (13)$$

where $P(k|k-1) = E [e(k|k-1)e(k|k-1)^T]$ is the covariance matrix of the prediction error $e(k|k-1) = x(k) - \hat{x}(k|k-1)$.

Note that the LRK model, on which the Kalman filter is applied, represents the ensemble of intermediate basins for a specific configuration. The Kalman Filter algorithm will then assimilate the discharge at every considered gauging stations and simultaneously correct all the states of the system.

TESTING METHODOLOGY

Split sample testing

The model was tested using the split-sample test proposed by Klemes (1986). It consists in splitting the available time series into two sub-periods (here, P1: 1994-2001 and P2: 2001-2009, including a 1-year warm-up period). The identification procedure is performed on P1 and the validation on P2, and then the role of the two periods is reversed (calibration on P2 and validation on P1). Thus the model can be evaluated in validation mode on the whole time span.

Parameter identification was done using the Levenberg-Marquardt algorithm with the root mean square errors (RMSE) criterion on the discharge at the outlet. For the application on the Seine River basin, the identification is done for each tested configuration and for each intermediate or upstream basin independently, i.e. using observed upstream flow data in the case of an intermediate basin. This procedure has two main advantages: it prevents the propagation of modelling errors from the upstream to the downstream basins and it reduces the amount of parameters to be identified simultaneously. Model parameters were calibrated in simulation mode, i.e. without considering the assimilation scheme.

In validation, the model was applied in hindcasting mode, i.e. applying the model retrospectively at each time-step of the test period as if it was in real-time conditions. At each time-step, the system states are corrected using the new observations and the data assimilation scheme, and the model produces forecasts with the corrected states as initial conditions and scenarios of future rainfall as inputs. In an operational context, such scenarios are produced by weather forecast centres (e.g., MétéoFrance or the European Centre for Medium-range Weather Forecast, ECMWF). The quality of rainfall forecasts impacts the performances of flood forecasting models (see e.g., Cloke and Pappenberger 2009; Rabuffetti and Barbero 2005), but archives of past precipitation forecasts were not available and this question is out of the scope of this study. Here two simple forecasting scenarios were considered:

- Scenario P0 : zero future rainfall. This assumption represents an unfavourable sce-

nario in the context of flood forecasting but is the default option when no rainfall forecast is available.

- Scenario PP : *perfect* rainfall forecast corresponding to the *a posteriori* observations, i.e. the best available precipitation estimates.

Using the rainfall observed *a posteriori* as future scenario (PP) put the model in an ideal situation, i.e. with limited rainfall uncertainty, and gives very optimistic results on model performances compared to what it could be in real time. The P0 scenario is investigated to test the model in more difficult conditions, i.e. without any information on future rainfall, which is very pessimistic in the perspective of flood forecasting.

Evaluation criteria

As mentioned in the introduction section, three days of anticipation are requested in Paris to organize evacuation and rescue. The maximum lead time considered in this study was then 72 hours. Model performance analysis was evaluated using the RMSE between discharge observations and forecasts for different lead times L ranging from 1 to 72 hours:

$$\text{RMSE}(L) = \sqrt{\frac{1}{N_{\mathcal{F}}} \sum_{k+L \in \mathcal{F}} (Q_{obs}(k+L) - Q_{for}(k+L))^2} \quad (14)$$

where $Q_{obs}(k+L)$ and $Q_{for}(k+L)$ are the observed and forecast downstream discharges at time $k+L$, respectively, \mathcal{F} a set of time-steps of the test period when the model is evaluated and $N_{\mathcal{F}}$ the number of time-steps in \mathcal{F} . The computation was restricted to high-flow periods, here defined above a flood threshold equal to the 0.95 (non-exceedance) quantile Q_{95} of the flow duration curve:

$$\mathcal{F} = \{k | Q_{obs}(k) \geq Q_{95}\} \quad (15)$$

This flood threshold was chosen to uniformly treat all the stations. As a benchmark, we also computed the RMSE for the persistence model which assumes that forecast flows equal the observed flow at the time of issuing the forecast, i.e. flows remain unchanged in

the future. Note that other evaluation scores and criteria could have been considered, some measuring the modelling error at each time step and others measuring the capacity of the model to forecast the exceedance of a threshold. RMSE was chosen here because it integrates all the model errors over the time series, and put more weight (due to the square) to large errors that generally occur in high flow conditions. Also this is a very common criterion, which appears in widely used efficiency measures (like the Nash-Sutcliffe criterion or the persistence index).

To compare the performance of the various tested configurations, we introduced another criterion called Forecast Performance Index (FPI), that accounts for the improvement for leading horizons ranging from 1 to 72 h, compared to the persistence model:

$$\text{FPI} = \frac{1}{72} \sum_{L=1}^{72} \frac{1}{2} \left(\frac{\text{RMSE}^{PP}(L)}{\text{RMSE}^{PERS}(L)} + \frac{\text{RMSE}^{P0}(L)}{\text{RMSE}^{PERS}(L)} \right) \quad (16)$$

where RMSE^{PP} and RMSE^{P0} are the RMSE criteria for the PP and P0 scenarios for a given configuration, while RMSE^{PERS} corresponds to the persistence model. FPI is lower than 1 for a configuration that has better performances than the persistence model. Note that FPI only gives an overview of the overall performances, since it does not inform on performance differences between small and large lead times or between P0 and PP scenarios.

The objective function used for parameter calibration is the RMSE computed in simulation mode. To give an overview of model performance in simulation mode, an evaluation of model performance was also made in calibration and validation using the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe 1970).

Tested configurations

The main objective was to evaluate model sensitivity at the outlet (Paris-Austerlitz station) using various upstream configurations. The simplest configuration is to apply the model in a lumped mode, i.e. only considering the Paris-Austerlitz station. This configuration serves as a reference and is noted A0 in the following.

The performance analysis was done in two steps. First, only one station upstream of Paris-Austerlitz among the four closest stations (Bazoches-lès-Bray, Courlon-sur-Yonne, Longing at Episy, La Ferté-sous-Jouarre) is considered. The four corresponding configurations are noted A1 to A4 respectively. Two additional configurations were considered by including the four upstream stations within the Yonne basin (Dornecy, Arcy-sur-Cure, Dissangis and Aisy-sur-Armanon, noted A2') or the second upstream station on the Marne basin (Châlon-sur-Marne, noted A4'). This aimed at evaluating the usefulness of these upstream stations. It can be expected that discharge forecasts be improved at the Courlon-sur-Yonne station in the case of A2' and at the Ferté-sous-Jouarre station in the case of A4'. A summary of the six A configurations is given in Table 2.

The four configurations with a single upstream station were ranked by increasing performance for flow forecasting at Paris-Austerlitz station using the RMSE(L) criterion. Then we tested several configurations with an increasing number of upstream stations, first including the station that provided the most successful results among the A1-A4 configurations, then the second most successful station, etc. These configurations will be noted B1 to B4.

RESULTS AND DISCUSSION

Results with only one upstream station

Calibration/validation results in simulation mode

For each configuration A1 to A4, models were first evaluated in simulation mode by applying the split-sample test previously described. Fig. 7 shows the NSE values at Paris-Austerlitz station obtained in simulation mode for periods P1 and P2 and for both calibration and validation steps. First, all the NSE values are larger than 0.70 which is a quite good score (see e.g., Chiew and McMahon 1993). Second, adding one upstream station in the model always improved the NSE at Paris-Austerlitz station, except for the A1 configuration (i.e. considering the Marne basin) on the P1 period (for identification as well as for the validation) which is due to the fact that no data are available before 1999 at Bazoches-lès-Bray (station

considered in A1). Moreover, for each period, NSE values are very close for identification and validation, which is an indication of the consistency of the optimised parameter sets.

In spite of good NSE values for the different configurations, assimilating discharge data proves of high added value. Indeed, whatever the configuration, the NSE is very close to 1 when the Kalman filter is applied. This shows that assimilating discharge data allows to efficiently correct the model state that is then used as the initial state in the forecast step.

Comparison of results in forecasting mode

Forecast performances of each configuration A1 to A4 (including A2' and A4') are compared to those of A0 for both scenarios PP (perfect rainfall foreknowledge) and P0 (zero future rainfall). Figure 8 shows the evolution of RMSE(L) averaged over P1 and P2 periods in validation mode. The performance of the persistence model is also plotted.

First it can be mentioned that performance curves are almost similar between the PP and P0 scenarios up to a certain lead time (between ten and 20 hours in most cases). This corresponds to the time response of the catchment: before this limit, the catchment response only depends on the rainfall fallen before the time of issuing the forecast and is therefore insensitive to future rainfall hypothesis. Second, as expected, the performance with the P0 scenario is always worse than with the PP scenario and the longer the lead time, the larger the difference. Model performance with P0 scenario becomes even worse than for the persistence model for lead times longer than 50 hours, whatever the configuration.

The comparison between A0 and A1 (first line in Figure 8) shows that including the Bazoches-lès-Bray station (Seine) does not improve the performance of the downstream discharge forecast. As stated before, this is explained by the fact that no data are available at this station before 1999. Indeed, on the 1994-1998 period, the discharge estimated at this station is not corrected during the assimilation process, which impacts the results.

The A2 configuration (second line), in which the gauging station at the outlet of the Yonne basin is considered, presents significantly better results than A0. This means that the Yonne tributary provides a significant part of the variability of the Seine streamflow

and that the TGR model with the Kalman Filter is able to well reproduce the discharge at the outlet of the Yonne basin. Including the four stations within the Yonne basin (i.e. A2' configuration) leads to an notable improved forecast at Courlon-sur-Yonne (Figure 9) that results in a better prediction for Paris-Austerlitz (3rd line, Figure 8), namely for the P0 scenario.

The A0-A3 comparison (fourth line) shows a very small improvement for the P0 scenario and for large lead times when the downstream station of the Loing basin is taken into account. This is explained by the fact that this basin only makes a limited contribution to floods on the downstream part of the Seine (the Loing represents only 9 % of the Seine basin area).

Finally, forecast performances of A0 and A4 configurations (fifth line) are quite similar. This result seems surprising since the Marne basin is a quite large contributor to the downstream flow. The difficulty for the hydrological model to satisfactorily simulate the behaviour of this basin may partly explain this result. The integration of the Châlons-sur-Marne station (configuration A4') yielded substantially better results at the Ferté-sous-Jouarre (see Figure 10), hence indicating the benefit of upstream observations. However benefits at Paris-Austerlitz remained modest (last line in Figure 8). The quality of data at the Ferté-sous-Jouarre station is also potentially responsible for limited performances. Indeed poor data quality could lead to corrected system states far from the reality. This assumption is explored in the next section.

As expected, the lowest FPI value is obtained with configuration A2, which has the highest upstream catchment area and mean discharge (Table 1). On the other hand, configuration A3 which obtained the second performance results, has the lowest upstream catchment area and mean discharge. This shows that best performances are not necessarily obtained when considering the most contributing upstream basins, in terms of both catchment area and mean discharge.

Results with several upstream stations

Configurations

In this section, the successive inclusion of several upstream stations is considered. Given the previous results, the upstream stations were sorted by decreasing gain in forecast efficiency (FPI) and the following order was chosen to successively include stations in the semi-distributed model: (1) Courlon-sur-Yonne, (2) Episy, (3) Bazoches-lès-Bray and (4) La Ferté-sous-Jouarre. These four new configurations will be named B1 to B4 hereafter and will also be compared to the reference configuration, noted B0, representing the entire basin without upstream stations. Note that configurations B0 and B1 are identical to A0 and A2, respectively.

Calibration/validation results in simulation mode

As for configurations A, Fig. 11 shows the NSE efficiency at Paris-Austerlitz station for configurations B. For period P2, the NSE always increase when new stations are added, leading to very high values (up to 0.95). The only exception is for the validation step for configuration B4. This case is discussed in the next section. Concerning period P1, there is a loss of efficiency from B1 to B2, which is probably due to the lack of data at Bazoches-lès-Bray before 1999. Besides, the parameter values seem to be still consistent since identification and validation NSE values remain close for each period.

Comparison of results in forecasting mode

Each configuration (including n upstream stations) was compared with the previous one (including $n-1$ stations) through the RMSE(L) and FPI criteria to visualize the improvements due to the successive addition of gauging station flows for use in data assimilation. Results are presented in Figure 12.

The B0-B1 comparison (first line) is identical to the A0-A2 comparison, showing the results of including the station Courlon-sur-Yonne. The integration of the Loing basin only marginally improves performance (second line). The B2-B3 comparison (third line) also shows that the integration of the upstream station on the Seine only slightly improves per-

formances. Remember that for this station, the performance is further enhanced when one considers only the period after 1999 without any gap. Finally, as shown by comparing B3-B4 (fourth line), the inclusion of the station at La Ferté-sous-Jouarre (Marne) deteriorates the forecast performance, even when the station of Châlons-sur-Marne is integrated (configuration B4'). This result is quite surprising as one might expect *a priori* that the integration of additional data in the data assimilation algorithm should lead to a better simulation of the downstream discharge.

As stated previously, a possible cause of this degradation is the quality of data at La Ferté-sous-Jouarre station. Unfortunately, no information on the data quality is available that could support this assumption. As a workaround, we increased the value related to the observations at this station in the measurement noise covariance matrix (R), which is equivalent to considering that measurements from this station are more uncertain. By doing this, the Kalman Filter gives less weight to this observation in the update step. The resulting configuration is called B4R. Figure 13 compares the performances for configuration B4 and B4R. Increasing the measurement noise correlation of this station highly increased the forecast performances. The value of FPI decreased from 0.988 to 0.754. This result clearly supports the assumption of poor data quality. It also shows how information about data quality, when available, can be taken into account in the data assimilation procedure.

Finally, Figure 14 shows the FPI values for all the tested configurations and for the GRP model. These results show that configuration B3 provides the best performance according to the FPI criterion. The forecast performance comparison between this configuration and the GRP model (currently used in operational conditions by the FFC) is presented in Figure 15, whereas Figure 16 shows an example of discharge forecast with GRP and B3 during the major flood of 1995. The TGR semi-distributed model with configuration B3 clearly outperformed the GRP lumped model (FPI of 0.722 and 0.810, respectively), namely for small lead times and with the PP scenario.

CONCLUSION

The objective of this study was to investigate the sensitivity of flood forecasts to the spatial distribution of inputs using a semi-distributed hydrological model (TGR). The semi-distributed approach was chosen to test different degrees of spatialization. Here the spatial distribution was defined by a subset of available gauging stations and related intermediate basins. Forecast performances of different spatial distribution are analyzed and compared to a reference lumped model (GRP) currently used operationally by the regional flood forecasting centre.

It is shown that the TGR model provides better forecast performance than the GRP lumped model. This result was expected since, despite the linearization of the routing reservoir, the TGR model uses observed upstream flow data not used by the GRP model.

Nevertheless, results obtained when increasing the degree of spatialization shows that including more gauging stations in the model does not systematically improve its performance. In an operational context, the model used has to be robust and computationally efficient, and it could be of prior importance to find the optimal spatial distribution which is not necessarily the one using all the available observed data. In addition, the case of the Marne river shows that considering some particular stations could even deteriorate the forecast performances. For such stations, we showed that increasing the measurement noise covariance can highly improve the forecast performances.

Another important advantage of the semi-distributed approach is that it is possible to handle parts of the basin with large human influences like regulated dams, especially when little information on the management rules is available. Considering a station downstream of the dam is a way to isolate its influence. Some current work focuses on the integration of a simple reservoir module into the TGR model (Ficchi et al. 2013).

Finally, Lerat et al. (2012) showed that decomposing the inflows into localized and distributed lateral inflows improves the discharge simulation inside the basin (anywhere in the main stream). Such a decomposition is possible with the routing model developed in this

study. Potential applications thus include the reconstruction and forecasting of the discharge at ungauged sites and using it as input for a 2D flooding model in order to elaborate prevention plans.

ACKNOWLEDGMENT

The authors thank Météo-France for providing the meteorological data used in this study and the Seine-Moyenne-Yonne-Loing flood forecasting centre (SPC SMYL) at DRIEE Ile-de-France for providing hydrological data, financial support and feedback on the results of the study. Finally, we are grateful to the anonymous reviewers for their comments that helped to improve the article.

SUPPLEMENTAL DATA

Appendixes S1-S3 are available online in the ASCE Library (www.ascelibrary.org).

REFERENCES

- Amengual, A., Diomedes, T., Marsigli, C., Martín, A., Morgillo, A., Romero, R., Papetti, P., and Alonso, S. (2008). “A hydrometeorological model intercomparison as a tool to quantify the forecast uncertainty in a medium size basin.” *Natural Hazards and Earth System Science*, 8(4), 819–838.
- Berthet, L., Andréassian, V., Perrin, C., and Javelle, P. (2009). “How crucial is it to account for the Antecedent Moisture Conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments.” *Hydrology and Earth System Sciences Discussions*, 6(2), 1707–1736.
- Billen, G., Silvestre, M., Barles, S., Mouchel, J.-M., Garnier, J., Curie, F., and Boët, P. (2009). “Le bassin de la Seine. Découvrir les fonctions et les services rendus par le système Seine.” *Agence de l’Eau Seine Normandie, Programme PIREN-Seine*, Fascicule, 51 p.
- Brown, R. and Hwang, P. (1992). *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley and Sons.

- Chiew, F. and Mc Mahon, T. (1993). "Assessing the adequacy of catchment streamflow yield estimates." *Australian Journal of Soil Research*, 31(5), 665.
- Cloke, H. L. and Pappenberger, F. (2009). "Ensemble flood forecasting: A review." *Journal of Hydrology*, 375, 613–626.
- Cole, S. J. and Moore, R. J. (2009). "Distributed hydrological modelling using weather radar in gauged and ungauged basins." *Advances in Water Resources*, 32(7), 1107–1120.
- Duan, Q., Sorooshian, S. and Gupta, V. (1992). "Effective and efficient global optimization for conceptual rainfall-runoff models.." *Water resources research*, 28(4), 1015–1031.
- Fan, P. and Li, J. (2006). "Diffusive wave solutions for open channel flows with uniform and concentrated lateral inflow." *Advances in Water Resources*, 29(7), 1000–1019.
- Ficchi, A., Raso, L., Jay-Allemand, M., Dorchies, D., Malaterre, P.-O., Pianosi, F., and Van Overloop, P.-J. (2013). "A centralized real-time controller for the reservoir's management on the Seine River using ensemble weather forecasting." *EGU General Assembly Conference Abstracts*, 15, 10420.
- Kalman, R. (1960). "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering*, 82-D, 35–45.
- Kealey, T. (2007). "Rising levels of discontent." *The Times*, August 20, Science Notebook.
- Klemes, V. (1986). "Operational testing of hydrological simulation models." *Hydrological Sciences Journal*, 31(1), 13–24.
- Lacaze, Y., Chesneau, S., Raimbault, E., Piotte, O., Silva, J.-P., Perrin, C., Andreassian, V., Berthet, L., Coron, L., and Fortier-Filion, T.-C. (2011). "Les modèles de prévision opérationnels d'aujourd'hui auraient-ils été fiables sur la crue de 1910 ? Analyse rétrospective critique sur une base de données de 1910 (Would the currently operational models have been reliable to forecast the 1910 flood event? Critical analysis using a historical database)." *La Houille Blanche*, (1), 22–29.
- Lerat, J., Perrin, C., Andréassian, V., Loumagne, C., and Ribstein, P. (2012). "Towards robust methods to couple lumped rainfall-runoff models and hydraulic models: A sensitivity

- analysis on the Illinois River.” *Journal of Hydrology*, 418-419, 123–135.
- Liu, Y., Weerts, A. H., Clark, M., Franssen, H. J. H., Kumar, S., Moradkhani, H., Seo, D. J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P. (2012). “Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities.” *Hydrology and Earth System Sciences*, 16(10), 3863–3887.
- McMillan, H. K., Hreinsson, E. O., Clark, M. P., Singh, S. K., Zammit, C., and Uddstrom, M. J. (2013). “Operational hydrological data assimilation with the recursive ensemble Kalman filter.” *Hydrology and Earth System Sciences*, 17(1), 21–38.
- Mendoza, P. A., McPhee, J., and Vargas, X. (2012). “Uncertainty in flood forecasting: A distributed modeling approach in a sparse data catchment.” *Water Resources Research*, 48(9), W09532.
- Munier, S. (2009). “Modélisation intégrée des écoulements pour la gestion en temps réel d’un bassin versant anthropisé (Integrated flow modeling for real time management of an anthropized basin).” Ph.D. thesis, AgroParisTech, Montpellier,
- Munier, S., Litrico, X., Belaud, G., and Malaterre, P.-O. (2008). “Distributed approximation of open-channel flow routing accounting for backwater effects.” *Advances in Water Resources*, 31(12), 1590–1602.
- Nash, J. and Sutcliffe, J. (1970). “River flow forecasting through conceptual models part I ? A discussion of principles.” *Journal of Hydrology*, 10(3), 282–290.
- Nester, T., Kirnbauer, R., Parajka, J., and Blöschl, G. (2012). “Evaluating the snow component of a flood forecasting model.” *Hydrology Research*, 43(6), 762.
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C. (2005). “Which potential evapotranspiration input for a lumped rainfall?runoff model?.” *Journal of Hydrology*, 303(1), 290–306.
- Perumal, M., Sahoo, B., Moramarco, T., and Barbetta, S. (2009). “Multilinear Muskingum Method for Stage-Hydrograph Routing in Compound Channels.” *Journal of Hydrologic*

- 663 *Engineering*, 14(7), 663–670.
- 664 Rabuffetti, D. and Barbero, S. (2005). “Operational hydro-meteorological warning and real-
665 time flood forecasting: the Piemonte Region case study.” *Hydrol. Earth Syst. Sci.*, 9,
666 457–466.
- 667 Refsgaard, J. C. (1997). “Validation and intercomparison of different updating procedures
668 for real-time forecasting.” *Nordic Hydrology*, 28(2), 65–84.
- 669 Smith, M. B., Koren, V., Zhang, Z., Zhang, Y., Reed, S. M., Cui, Z., Moreda, F., Cosgrove,
670 B. A., Mizukami, N., and Anderson, E. A. (2012). “Results of the DMIP 2 Oklahoma
671 experiments.” *Journal of Hydrology*, 418-419, 17–48.
- 672 Todini, E. (2007). “Hydrological catchment modelling: past, present and future.” *Hydrology
673 and Earth System Sciences*, 11(1), 468–482.
- 674 Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M. (2010).
675 “A 50-year high-resolution atmospheric reanalysis over France with the Safran system.”
676 *International Journal of Climatology*, 30(11), 1627–1644.
- 677 Viennot, P., Ducharne, A., Habets, F., Lamy, F., and Ledoux, E. (2009). “Hydrogéologie
678 du bassin de la Seine.” *Agence de l’Eau Seine-Normandie, Programme PIREN-Seine*,
679 Fascicule, 55 p.
- 680 Villion, G. (1997). “Rôle des lacs-réservoirs amont : les grands lacs de Seine.” *La Houille
681 Blanche*, 8, 51–56.
- 682 Welles, E., Sorooshian, S., Carter, G., and Olsen, B. (2007). “Hydrologic verification. A call
683 for action and collaboration.” *Bull. Amer. Meteor. Soc.*, 88 (4), 503–511.
- 684 Young, P. C. (1974). “Recursive approaches to time series analysis.” *Bull. Inst. Math. Appl.*,
685 10, 209–224.
- 686 Young, P. C. (2002). “Advances in real-time flood forecasting..” *Philosophical transactions.
687 Series A, Mathematical, physical, and engineering sciences*, 360(1796), 1433–50.

688	List of Tables	
689	1	List of the flow gauging stations. The mean discharge and the percentage of
690		missing values are given for the 1995-2009 period. 29
691	2	List of tested configurations A0 to A4, and gauging stations considered. . . . 30

TABLE 1. List of the flow gauging stations. The mean discharge and the percentage of missing values are given for the 1995-2009 period.

Station code	Station name	River	Area (km ²)	Mean discharge (m ³ /s)	Missing values (%)
H1940020	Bazoches-lès-Bray	Seine	10100	85.2	35.6
H2051010	Dornecy	Yonne	754	10.0	7.8
H2182010	Arcy-sur-Cure	Cure	1182	16.4	9.6
H2332020	Dissangis	Serein	636	4.5	6.1
H2452020	Aisy-sur-Armanon	Armanon	1350	13.1	7.5
H2721010	Courlon-sur-Yonne	Yonne	10700	99.8	7.3
H3621010	Episy	Loing	3900	20.5	7.4
H5201010	Châlon-sur-Marne	Marne	6280	75.7	7.0
H5321010	La Ferté-sous-Jouarre	Marne	8818	97.9	12.8
H5920010	Paris-Austerlitz	Seine	43800	343.7	11.1

TABLE 2. List of tested configurations A0 to A4, and gauging stations considered.

Configuration	Station names
A0	Paris-Austerlitz
A1	Paris-Austerlitz, Bazoches-lès-Bray
A2	Paris-Austerlitz, Courlon-sur-Yonne
A2'	Paris-Austerlitz, Courlon-sur-Yonne, Dornecy, Arcy-sur-Cure, Dissangis and Aisy-sur-Armanon
A3	Paris-Austerlitz, Episy
A4	Paris-Austerlitz, La Ferté-sous-Jouarre
A4'	Paris-Austerlitz, La Ferté-sous-Jouarre, Châlon-sur-Marne

List of Figures

1	The Seine River basin at Paris (43,800 km ²), and the networks of raingauges and flow gauging stations used in the study. The station details corresponding to the codes shown on the map are given in Table 1.	33
2	Schematic representation of the hydrology-routing coupled model (PE and P are the potential evapotranspiration and rainfall inputs to the model, Q stands for streamflow). The drained area curve is used for the spatial discretization inside the intermediate basin.	34
3	(a) Example of hydrological discretization based on the gauging station network and (b) corresponding modelling scheme.	35
4	Structure of the hydrological model.	36
5	Flow routing model scheme.	37
6	The TGR coupled model scheme for one intermediate basin.	38
7	Nash-Sutcliffe efficiency obtained at Paris-Austerlitz station in simulation mode for A0 to A4 configuration in identification and validation for the two test periods P1 and P2.	39
8	Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the Paris-Austerlitz station for the six configurations (see Table 2) using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours. Two configurations are compared in each case, illustrated on the left and right hand sides of the graph.	40
9	Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the Courlon-sur-Yonne station for the A2 and A2' configurations using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours.	41

717	10	Forecast performance obtained in validation by the TGR model and the per-	
718		sistence model (PERS) at the La Ferté-sous-Jouarre station for the A4 and	
719		A4' configurations using PP and P0 rainfall scenarios. Lead times range from	
720		1 to 72 hours.	42
721	11	Nash-Sutcliffe efficiency obtained at Paris-Austerlitz station in simulation	
722		mode for B0 to B4 configuration in identification and validation for the two	
723		test periods P1 and P2.	43
724	12	Forecast performance obtained in validation by the TGR model and the per-	
725		sistence model (PERS) at the Paris-Austerlitz station for the B configurations	
726		using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours. Two	
727		configurations are compared in each case, illustrated on the left and right hand	
728		sides of the graph.	44
729	13	Forecast performance obtained in validation by the TGR model with configu-	
730		rations B4 and B4R, and the persistence model (PERS) at the Paris-Austerlitz	
731		station using PP and P0 rainfall scenarios. Lead times range from 1 to 72	
732		hours.	45
733	14	Forecast Performance Index for all the tested configurations and for the GRP	
734		model.	46
735	15	Forecast performance obtained in validation by the TGR model (configuration	
736		B3), the GRP model and the persistence model (PERS) at the Paris-Austerlitz	
737		station using PP and P0 rainfall scenarios. Lead times range from 1 to 72	
738		hours.	47
739	16	Example of discharge forecast by the TGR model (configuration B3) and the	
740		GRP operational model during the 1995 main flood event with a 72-hour lead	
741		time and for PP and P0 scenarios.	48

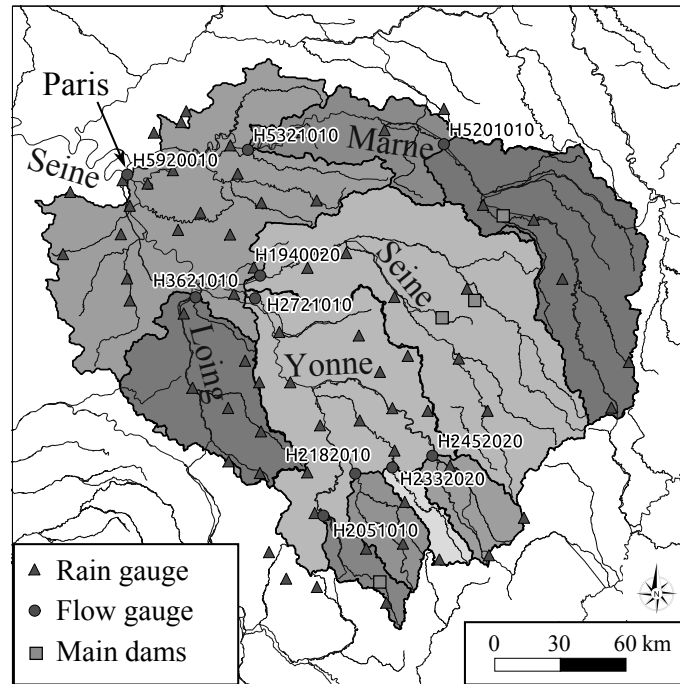


FIG. 1. The Seine River basin at Paris (43,800 km²), and the networks of raingauges and flow gauging stations used in the study. The station details corresponding to the codes shown on the map are given in Table 1.

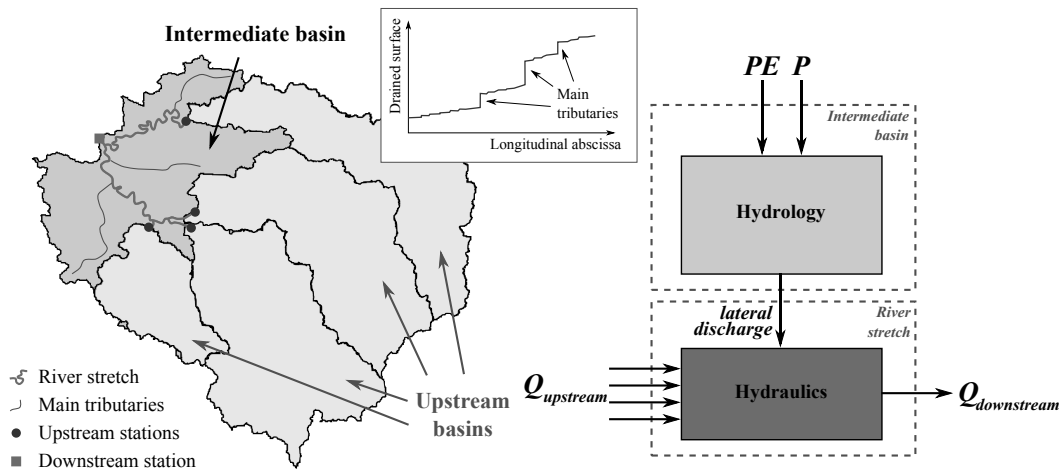


FIG. 2. Schematic representation of the hydrology-routing coupled model (PE and P are the potential evapotranspiration and rainfall inputs to the model, Q stands for streamflow). The drained area curve is used for the spatial discretization inside the intermediate basin.

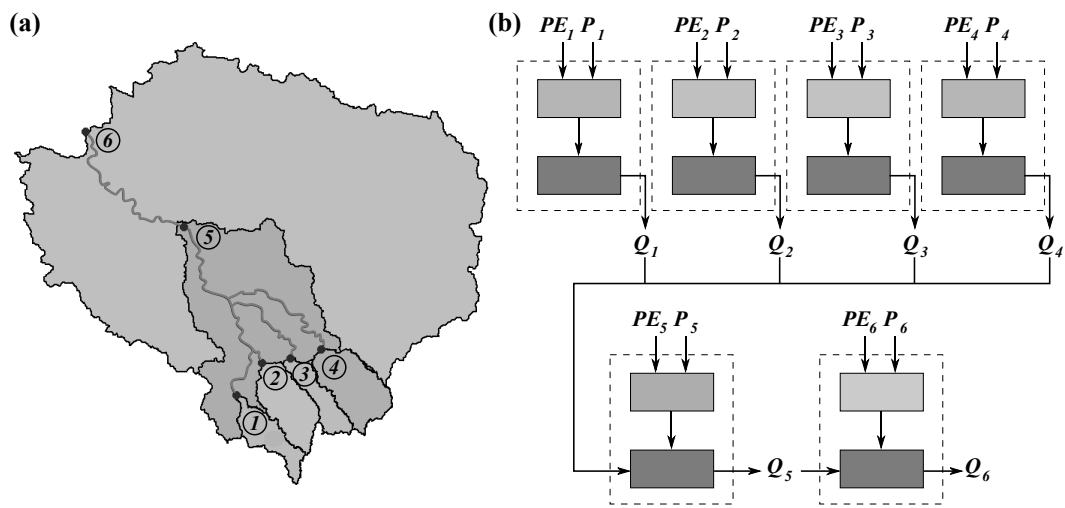


FIG. 3. (a) Example of hydrological discretization based on the gauging station network and (b) corresponding modelling scheme.

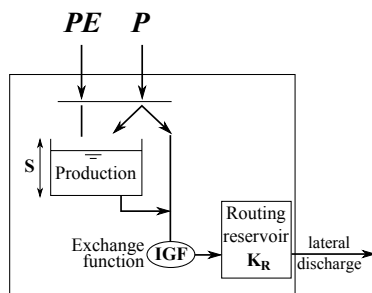


FIG. 4. Structure of the hydrological model.

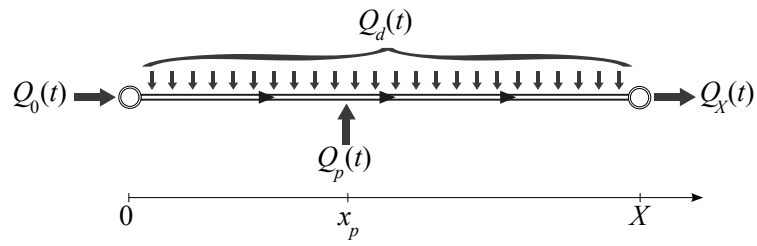


FIG. 5. Flow routing model scheme.

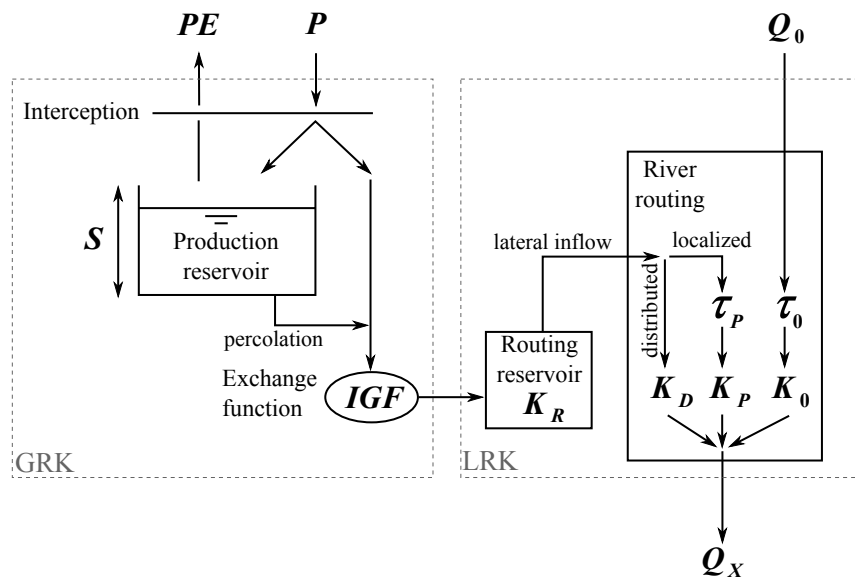


FIG. 6. The TGR coupled model scheme for one intermediate basin.

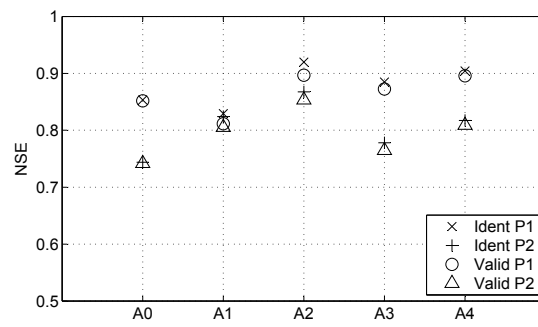


FIG. 7. Nash-Sutcliffe efficiency obtained at Paris-Austerlitz station in simulation mode for A0 to A4 configuration in identification and validation for the two test periods P1 and P2.

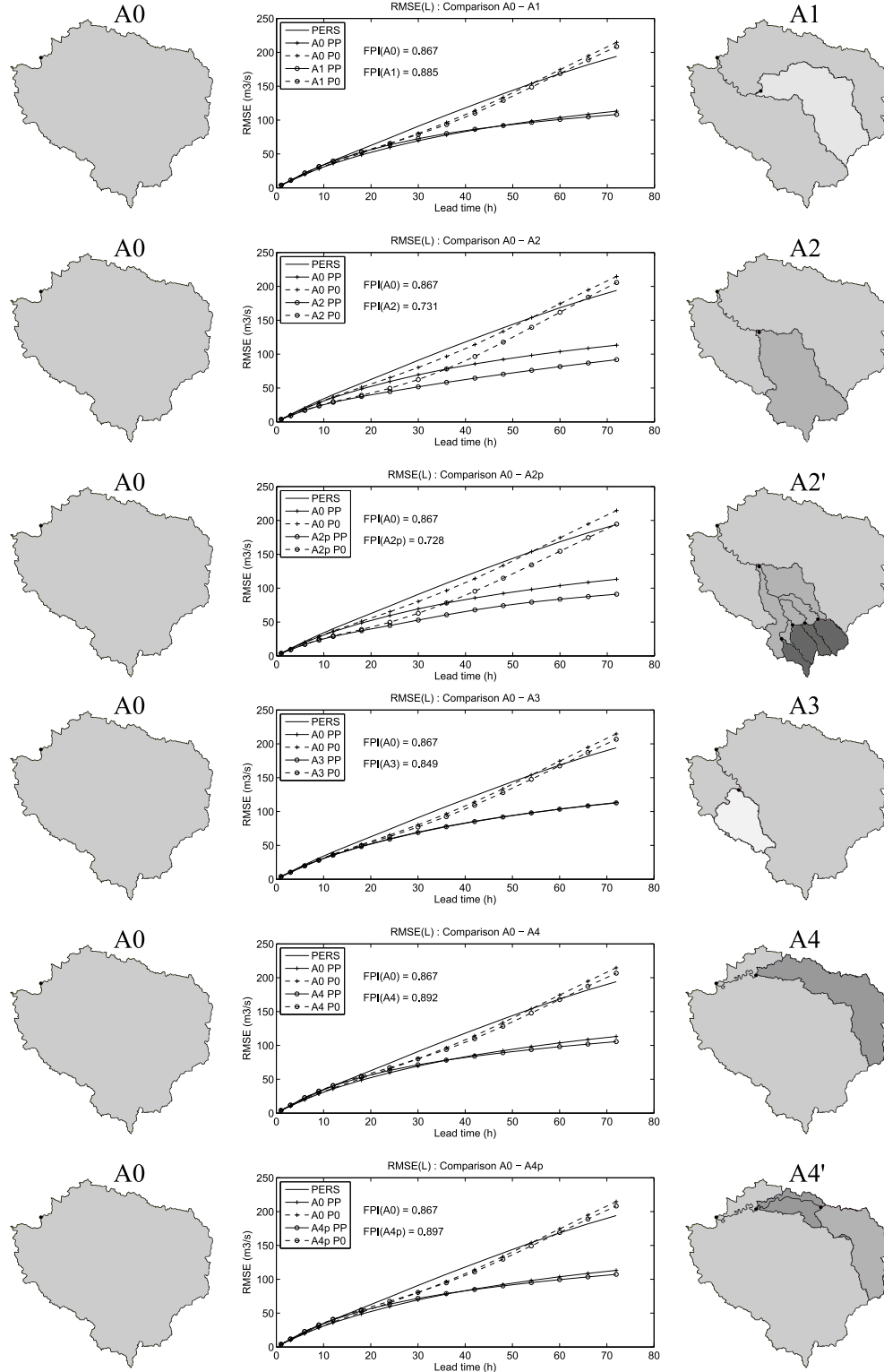


FIG. 8. Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the Paris-Austerlitz station for the six configurations (see Table 2) using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours. Two configurations are compared in each case, illustrated on the left and right hand sides of the graph.

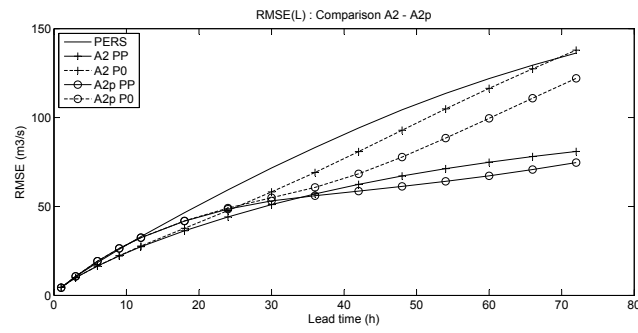


FIG. 9. Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the Courlon-sur-Yonne station for the A2 and A2' configurations using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours.

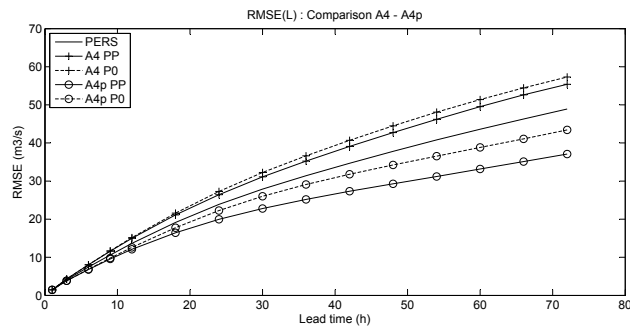


FIG. 10. Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the La Ferté-sous-Jouarre station for the A4 and A4' configurations using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours.

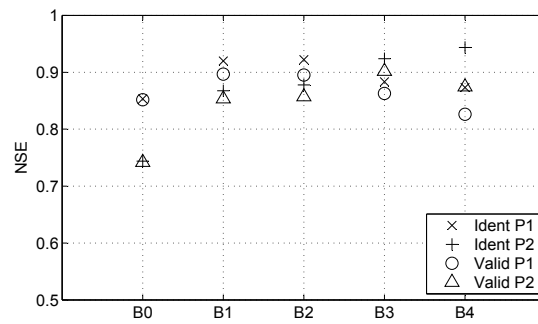


FIG. 11. Nash-Sutcliffe efficiency obtained at Paris-Austerlitz station in simulation mode for B0 to B4 configuration in identification and validation for the two test periods P1 and P2.

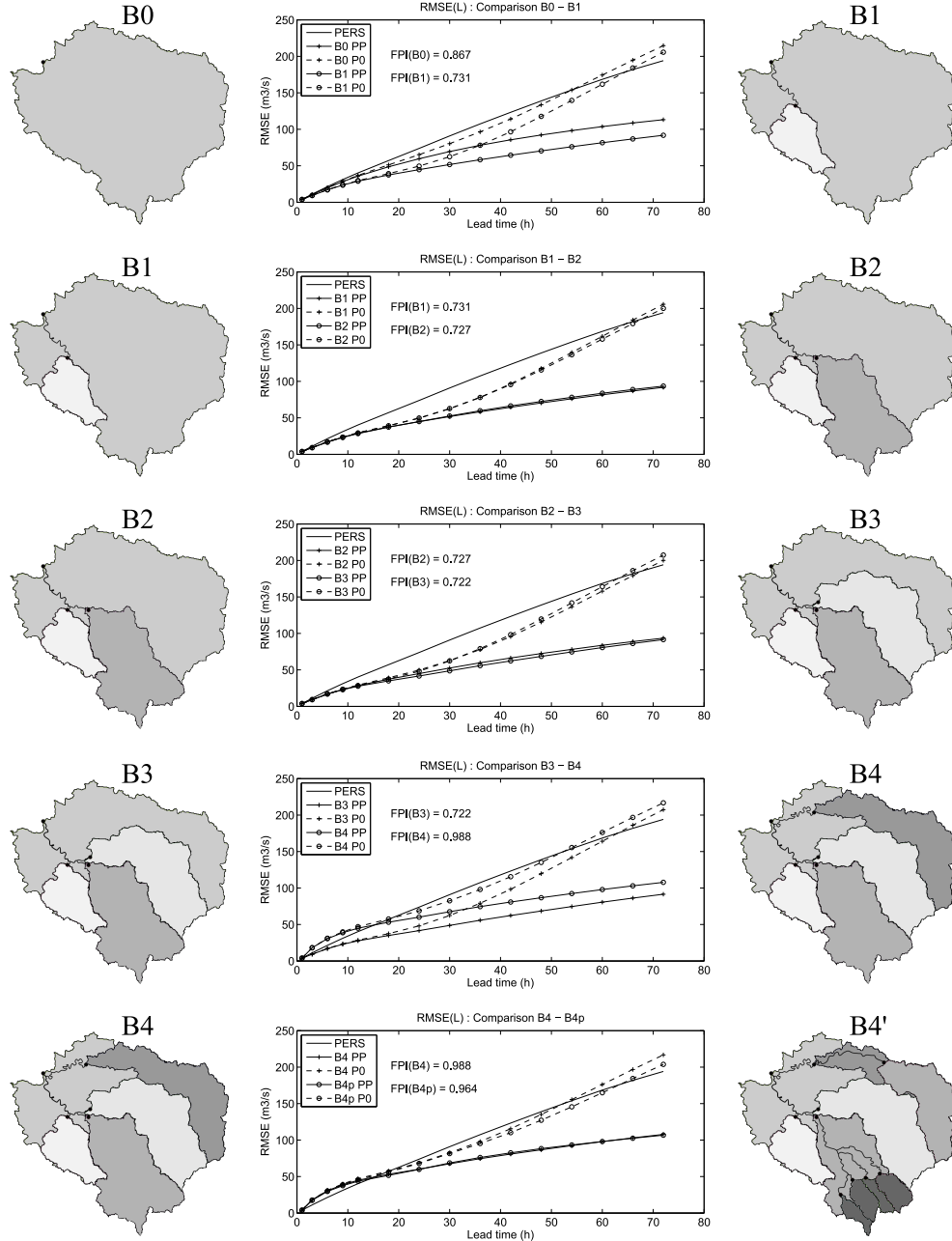


FIG. 12. Forecast performance obtained in validation by the TGR model and the persistence model (PERS) at the Paris-Austerlitz station for the B configurations using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours. Two configurations are compared in each case, illustrated on the left and right hand sides of the graph.

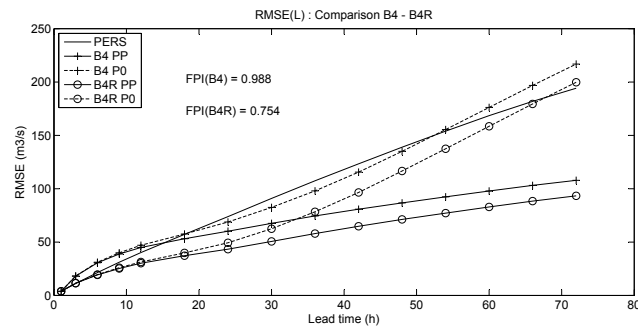


FIG. 13. Forecast performance obtained in validation by the TGR model with configurations B4 and B4R, and the persistence model (PERS) at the Paris-Austerlitz station using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours.

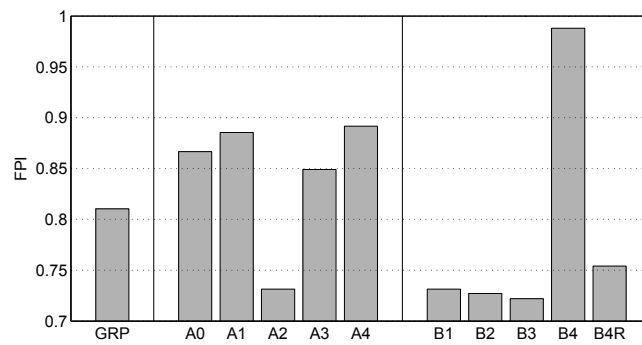


FIG. 14. Forecast Performance Index for all the tested configurations and for the GRP model.

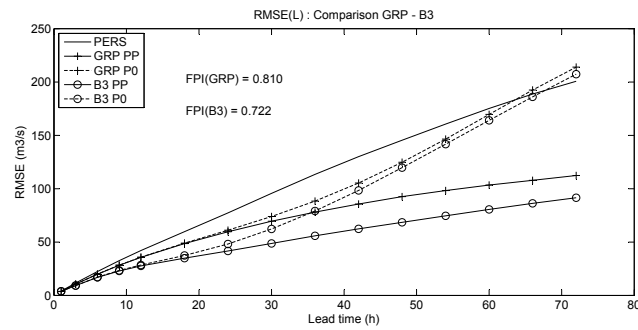


FIG. 15. Forecast performance obtained in validation by the TGR model (configuration B3), the GRP model and the persistence model (PERS) at the Paris-Austerlitz station using PP and P0 rainfall scenarios. Lead times range from 1 to 72 hours.

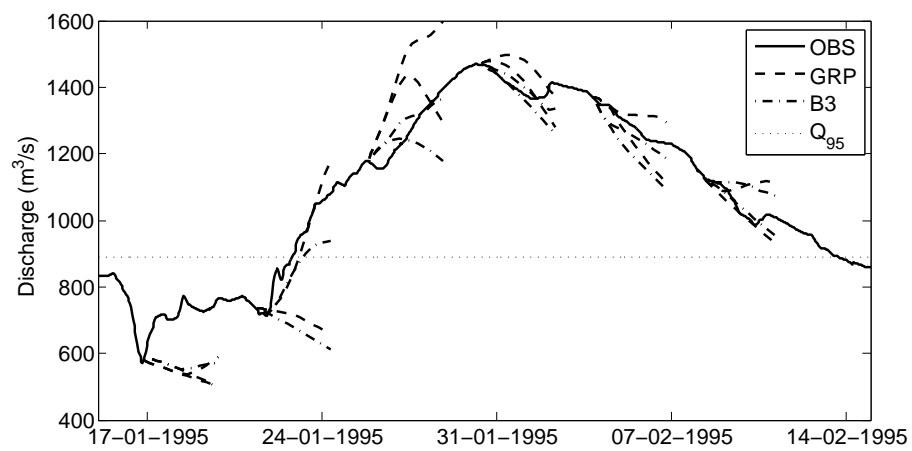


FIG. 16. Example of discharge forecast by the TGR model (configuration B3) and the GRP operational model during the 1995 main flood event with a 72-hour lead time and for PP and P0 scenarios.